The Application of the Yo-Yo Intermittent Endurance Level 2 Test to Elite Female Soccer Populations


1Department of Sport and Exercise, University of Sunderland, Sunderland, UK, 2Department of Exercise and Sport Sciences, University of Copenhagen, Copenhagen, Denmark, 3Olympique Lyonnais Football Club, Lyon, France, 4Sport and Health Sciences, College of Life and Environmental Sciences, St. Luke’s Campus, University of Exeter, Exeter, UK, 5The Football Association, England National Women’s Team, London, UK, 6Medical Department, FC Cartagena, Cartagena, Spain, 7Institute of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark.

Accepted for publication 23 April 2012

The aim of this study was to evaluate the application of the Yo-Yo intermittent endurance test level 2 (Yo-Yo IE2) to elite female soccer populations. Elite senior (n = 92), youth (n = 42), domestic (n = 46) and sub-elite female soccer players (n = 19) carried out the Yo-Yo IE2 test on numerous occasions across the season. Test–retest coefficient of variation (CV) in Yo-Yo IE2 test performance in domestic female players was 4.5%. Elite senior female players’ Yo-Yo IE2 test performances were better (P < 0.01) than youth, domestic and sub-elite players (mean ± standard deviation; 1774 ± 532 vs 1490 ± 447, 1261 ± 449, and 994 ± 373 m). For elite senior female players, wide midfielders (2057 ± 550 m) had a higher Yo-Yo IE2 test performance (P < 0.05) than central defenders (1588 ± 534 m) and attackers (1516 ± 401 m), but not central midfielders (1764 ± 473 m) or full-backs (1964 ± 522 m). Large correlations were observed between Yo-Yo IE2 test performance and the total and high-intensity distance covered (r = 0.55; P < 0.05) during elite senior soccer matches (r = 0.70; P < 0.01). A large correlation was also obtained between Yo-Yo IE2 test performance and VO₂max (r = 0.68; P < 0.01). Performances in the Yo-Yo IE2 test were greater (P < 0.05) in the middle and the end of the season compared with the preparation period for elite youth female players (1767 ± 539 and 1742 ± 503 vs 1564 ± 504 m) and in elite senior female players, Yo-Yo IE2 test performance increased by 14% (P < 0.01) after completing 4 weeks of intense training prior to the International Federation of Association Football Women’s World Cup Finals (2049 ± 283 vs 1803 ± 342 m). The data demonstrate that the Yo-Yo IE2 test is reproducible and is an indicator of the match-specific physical capacity of female soccer players. Furthermore, the Yo-Yo IE2 test illustrates sensitivity by differentiating intermittent exercise performance of female players in various competitive levels, stages of the season and playing positions.

In soccer, the mode of exercise is intermittent and performance is related to the players’ ability to perform repeated bouts of high-intensity exercise interspersed by recovery periods (Mohr et al., 2005; Bangsbo et al., 2006; Bradley et al., 2009; Krustup et al., 2010). Studies indicate that the relative physiological loadings during match play are similar across gender, suggesting that the aerobic system is heavily taxed throughout matches and particularly during intense periods (Ekblom, 1986; Bangsbo, 1994; Krustup et al., 2003, 2005, 2010; Stolen et al., 2005). However, female players seem to possess a lower physical capacity than male players across a range of aerobic and anaerobic fitness tests (Rhodes & Mosher, 1992; Tamer et al., 1997; Krustup et al., 2005, 2010; Stolen et al., 2005; Mujika et al., 2009). Thus, it is not surprising that studies have reported that high-intensity running distances covered in elite female matches are ~30% lower than their male counterparts of a similar competitive level and age (Krustup et al., 2005; Mohr et al., 2008; Andersson et al., 2010). However, limited research currently exists on the differences in physical capacity between elite female players competing at different levels of competition.

Both laboratory and field test modalities are currently used in elite female soccer in order to evaluate physical performance and training status. These include treadmill tests for the determination of maximal oxygen uptake (VO₂max) and various indices of lactate threshold, in addition to free running tests performed in the field (Krustup et al., 2005, 2010). To be qualified for evaluating elite female soccer players, tests must exhibit robust test–retest reproducibility and various forms of validity (concurrent and construct validity). The reproducibility of performance tests must be high...
and ultimately exhibit minimal changes from day to day (Hopkins, 2000). Previously, soccer players have demonstrated varying degrees of test–retest reproducibility on field-based performance tests possibly because of the population studied (Michalsik & Bangsbo, 1995; Krstrup et al., 2006b; Bangsbo et al., 2008; Bradley et al., 2011). Thus, it is imperative that population-specific reproducibility is obtained. Construct validity is the degree to which a test can easily discriminate between various subject populations such as elite and sub-elite female soccer players (Currell & Jeukendrup, 2008). Concurrent validity is demonstrated by a relationship between the test result and the physical performance of players during matches (Bangsbo & Lindquist, 1992; Krstrup et al., 2003). Thus, it would be prudent for applied sports scientists to establish both forms of validity before using various testing modalities to monitor the impact of training interventions on physical fitness parameters and match running performance. Although no exact measure of physical performance in elite female soccer matches exists, the total distance covered and particularly that performed at high intensity provide useful indicators of match running performance (Krustrup et al., 2005, 2010; Mohr et al., 2008; Andersson et al., 2010). Thus, the lack of association between test measures and match running performance in some studies could be attributed to the continuous nature of tests and their limited specificity to soccer (Krustrup et al., 2003; Castagna et al., 2005), although some authors have observed associations between continuous tests and match running performance (Rampinini et al., 2007; Buchheit et al., 2010). Intermittent free running protocols such as the Yo-Yo intermittent recovery tests (Yo-Yo IR1 and IR2) and the Yo-Yo intermittent endurance test level 2 (Yo-Yo IE2) are popular among practitioners in elite male and female soccer (Bangsbo, 1994; Bangsbo et al., 2006, 2008; Krstrup et al., 2003, 2005, 2006a; Bradley et al., 2011). Field-based tests are advantageous in the applied setting as a team can be tested frequently and rapidly at low cost (Svensson & Drust, 2005). Krstrup et al. (2005) found that the Yo-Yo IR1 test closely correlated with match running performance in elite female players. However, limited research exists concerning the Yo-Yo IE2 test, despite its frequent use by female players at all levels of competition (international senior/youth, domestic and sub-elite). Although limited data exist to support the Yo-Yo IE2 tests application to female soccer populations, some studies have still used the test to evaluate match-induced fatigue patterns after elite senior female soccer matches (Krustrup et al., 2010). Thus, a systematic study using a large sample of female players is needed to establish the Yo-Yo IE2 tests reproducibility, validity, and sensitivity. Thus, the aim of this study was to evaluate the application of the Yo-Yo IE2 test to elite female soccer populations.

**Table 1. Number of female soccer players participating in different subsections of the study**

<table>
<thead>
<tr>
<th>Population</th>
<th>Yo-Yo IE2</th>
<th>VO2max</th>
<th>Test–retest</th>
<th>Match analysis</th>
<th>Intra-seasonal changes</th>
<th>Heart rate response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elite senior</td>
<td>92</td>
<td>13</td>
<td>23</td>
<td>(10 + 13)</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>Elite youth</td>
<td>42</td>
<td></td>
<td></td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domestic</td>
<td>46</td>
<td>19</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sub-elite</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note that numbers within parentheses indicate the various subgroups of players within a similar competitive standard.

**Materials and methods**

**Subjects**

A total of 199 female soccer players participated in the study. The sample of elite senior female players (n = 92) consisted of two European national squads in addition to elite teams competing in the Union of European Football Associations (UEFA) Champions League (mean ± standard deviation; age, height, and body mass: 23 ± 2 years, 1.68 ± 0.07 m, and 61.9 ± 5.6 kg). The elite youth female population (n = 42) comprised of players within a European national U-20 squad (age, height, and body mass: 19 ± 1 years, 1.67 ± 0.06 m, and 59.9 ± 4.1 kg). The domestic (n = 46) and sub-elite female players (n = 19) were derived from the highest and lowest tiers of the game in two European countries (age, height and body mass: 22 ± 3 vs 23 ± 4 years, 1.69 ± 0.05 vs 1.65 ± 0.05 m and 60.9 ± 3.4 vs 58.2 ± 4.8 kg). Please refer to Table 1 for the number of players contained in each subsection of the study. The study conforms to the code of ethics of the World Medical Association and was approved by the appropriate institutional ethical committees.

**The Yo-Yo IE2 test**

All players (n = 199) completed the Yo-Yo IE2 test at the start of the season. The test lasts for 5–25 min and consists of repeated 20 m shuttle runs at progressively increasing speeds dictated by an audio bleep emitted from a CD player (Fig. 1). Between each shuttle, the players had a 5-s period of jogging around a marker placed 2.5 m behind the finishing line. Failure to achieve the shuttle run in time on two occasions resulted in termination of the test and the distance covered in the last complete successful shuttle was recorded as the test result. All testing sessions were performed indoors on an artificial grass surface within a 2 × 20 m running lane marked by cones. Before the test, all players carried out a warm-up period consisting of the first three running bouts of the Yo-Yo IE2 test followed by a period of lower-extremity stretching. All players had been familiarized with the Yo-Yo IE2 test and experimental procedures previously.

**Heart rate response during the Yo-Yo IE2 test**

In elite senior players’ (n = 28) heart rates were recorded by an automated monitor (Polar Electro Oy, Kempele, Finland) placed around the chest for continuous recordings throughout selected stages of the Yo-Yo IE2 test. Specifically, heart rates were collected before the warm-up, during the test (280 and 600 m), at exhaustion in addition to 1 and 2 min after the test. All heart rates before and after the Yo-Yo IE2 test were collected while players were standing.
were determined as the peak values reached in a 15-s period. All players had been familiarized to the criteria were met: (a) a change in maximal heart rate (HRmax) of each player ($n = 28$) was based on the competitive standard of the player. The maxima were only obtained from the domestic female players ($n = 19$) completed a laboratory treadmill test, which consisted of incremental stages until exhaustion. This test started at a running speed of 11–12 km/h for 2 min and progressed with a stepwise 1 km/h speed inclination every 30 s until exhaustion. The variation in the starting speed was based on the diurnal variation. The maximal heart rate (HR$_{\text{max}}$) of each player ($n = 47$) was determined from the peak values attained over a 5-s period during the latter stages of the test. Although HR$_{\text{max}}$ was collected for all players, VO$_{\text{max}}$ was only obtained from the domestic female players ($n = 19$). Pulmonary oxygen uptake was measured during the incremental treadmill test using a breath-by-breath gas analyzer (MedGraphics CPX/D, Saint Paul, MN, USA). This system was calibrated before testing with gases of known concentration and the tube flow meter was calibrated using a 3-L syringe. It was judged that the players had reached VO$_{\text{max}}$ when the following criteria were met: (a) a change in VO$_2$ < 2 mL/kg/min with increasing workload; (b) a respiratory exchange ratio > 1.15; and (c) a heart rate within 5 beats/min of the age-predicted maximum. Individual values for VO$_{\text{max}}$ were determined as the peak values reached in a 15-s period. All players had been familiarized to the laboratory incremental test previously.

**Test–retest reproducibility of the Yo-Yo IE2 test**

Domestic female players ($n = 27$) carried out the Yo-Yo IE2 test on two separate occasions interspersed by 7 days. All testing sessions were conducted at the same time of day to reduce the influence of diurnal variation.

**Laboratory incremental treadmill test**

Within 7 days of completing the Yo-Yo IE2 test, elite senior ($n = 28$) and domestic female players ($n = 19$) completed a laboratory treadmill test, which consisted of incremental stages until exhaustion. This test started at a running speed of 11–12 km/h for 2 min and progressed with a stepwise 1 km/h speed inclination every 30 s until exhaustion. The variation in the starting speed was based on the diurnal variation. The maximal heart rate (HR$_{\text{max}}$) of each player ($n = 47$) was determined from the peak values attained over a 5-s period during the latter stages of the test. Although HR$_{\text{max}}$ was collected for all players, VO$_{\text{max}}$ was only obtained from the domestic female players ($n = 19$). Pulmonary oxygen uptake was measured during the incremental treadmill test using a breath-by-breath gas analyzer (MedGraphics CPX/D, Saint Paul, MN, USA). This system was calibrated before testing with gases of known concentration and the tube flow meter was calibrated using a 3-L syringe. It was judged that the players had reached VO$_{\text{max}}$ when the following criteria were met: (a) a change in VO$_2$ < 2 mL/kg/min with increasing workload; (b) a respiratory exchange ratio > 1.15; and (c) a heart rate within 5 beats/min of the age-predicted maximum. Individual values for VO$_{\text{max}}$ were determined as the peak values reached in a 15-s period. All players had been familiarized to the laboratory incremental test previously.

**Intra-seasonal changes in Yo-Yo IE2 test performance**

Testing sessions were completed regularly throughout the season for all female soccer populations. However, for comparative purposes, a sample of elite youth female players ($n = 11$) were assessed using the Yo-Yo IE2 test in four phases of the international season leading up to the International Federation of Association Football (FIFA) U-20 Women’s World Cup (September, March, May, and June). For a group of elite senior players ($n = 10$), Yo-Yo IE2 test performance was monitored across the middle stage of the season (November, January, and March). While a second group of elite senior female players ($n = 13$) were assessed using the Yo-Yo IE2 test at two stages leading up to the FIFA Women’s World Cup (July and August). Finally, domestic female players ($n = 10$) also carried out the Yo-Yo IE2 test at the start and end of the season (September and April). The number of players completing all testing periods was influenced by a multitude of factors such as injury, illness, and transfers to other teams or failure to be selected for international duty.

**Match analysis**

Within 14 days of one of the Yo-Yo IE2 tests being performed, the time motion characteristics of elite female players ($n = 13$) were examined using a multiple-camera match analysis system (Amisco Pro®, Sport-Universal, Nice, France). Each player was tracked across a two-match period and the average provided an indication of match running performance. All cameras positioned at roof level were calibrated and synchronized to enhance the accuracy of the data. The movements of players over the course of the match were tracked at a sampling rate of 25 Hz. Match running performance was determined automatically by computerized analysis of the player movements and activities using match-analysis software (Amisco Viewer®, Sport-Universal Process, Nice, France). The accuracy and reproducibility of similar semi-automated systems has been described in more detail elsewhere (Di Salvo et al., 2006; Bradley et al., 2009). Total distance represented the overall distance covered during the match. Low-intensity distance covered during the match was classified as activity < 12 km/h. The traditional classification of high-intensity running was activity > 15 km/h (Krustrup et al., 2005; Mohr et al., 2008; Andersson et al., 2010). Players’ high-intensity activities were also coded into the following speed thresholds: 18–21, 21–23, 23–25, and 25–27 km/h. Sprinting activities were classified as activity > 27 km/h. The speeds for each category are similar to those employed in previous time-motion studies (Di Salvo et al., 2007; Carling et al., 2010). The low sample size included for the analyses was due to substitutions and/or injuries during matches but was comparable to other studies investigating...
match running performance and physical capacity relationships (Krustrup et al., 2003).

Statistical analysis
All data analyses were conducted using the appropriate statistical software [Statistical Package for the Social Sciences (SPSS) (SPSS Inc., Chicago, IL, USA)]. Descriptive statistics were calculated on each variable and z-scores were used to verify the normality of the data. Within-subject changes in heart rate before, during and after the Yo-Yo IE2 test were evaluated using a one-way analysis of variance (ANOVA) with repeated measures. Two separate one-way ANOVAs were used to evaluate differences between female soccer players at different competitive levels and playing positions for Yo-Yo IE2 test performance. In the event of a difference occurring, Tukey’s post-hoc tests were used to identify any localized effects. The coefficient of variation (CV) and typical error (TE) in absolute units were calculated to assess reproducibility of the domestic female players (Atkinson & Nevill, 1998; Hopkins et al., 2001). Furthermore, a threshold of 1.5–2.0 times the TE was used to indicate a worthwhile change in performance for domestic female players (Hopkins, 2000). The effect size (ES) was calculated to determine the magnitude of the treatment effect (Cohen, 1988). The magnitude of the ES was classified as trivial ($r < 0.2$), small ($r = 0.2–0.6$), moderate ($r = 0.6–1.2$), large ($r > 1.2–2.0$), and very large ($r > 2.0–4.0$) based on guidelines from Batterham and Hopkins (2006). Relationships between Yo-Yo IE2 test performance and selected variables (match running performance and VO$_{2\text{max}}$) were evaluated using Pearson’s product moment test. The magnitudes of the correlations were considered as trivial ($r < 0.1$), small ($r = 0.1–0.3$), moderate ($r = 0.3–0.5$), large ($r = 0.5–0.7$), very large ($r > 0.7–0.9$), nearly perfect ($r > 0.9$), and perfect ($r = 1.0$) in accordance with Hopkins et al. (2009). Statistical significance was set at $P < 0.05$. Values are presented as mean and standard deviations unless otherwise stated.

Results
Test–retest reproducibility of the Yo-Yo IE2 test
No difference was found for domestic female players when the Yo-Yo IE2 test was performed on two separate occasions separated by 7 days (1330 ± 392 vs 1384 ± 382 m; $n = 27$). The intra-individual difference between these two tests averaged 53 ± 94 m, with a CV and TE of 4.5% and 67 m, respectively (Fig. 2).

Heart rate response to the Yo-Yo IE2 test
For elite senior female players ($n = 28$), heart rate was 96 ± 19 bpm before the test and increased ($P < 0.01$) to 176 ± 7 and 184 ± 8 bpm after 280 and 600 m (ES range of 1.1–6.0), respectively. This corresponded to 90 ± 3 and 96 ± 2%HR$_{\text{max}}$. At exhaustion, heart rate was 192 ± 6 bpm or 98 ± 2%HR$_{\text{max}}$. During the recovery period, heart rate decreased from exhaustion ($P < 0.01$) to 145 ± 19 and 120 ± 14 bpm after 1 and 2 min (ES range of 6.0–6.6), respectively. This corresponded to 74 ± 9 and 61 ± 7%HR$_{\text{max}}$. A very large inverse relationship was observed between individual performances of the Yo-Yo IE2 test and %HR$_{\text{max}}$ reached at 4 min ($r = -0.74$; $P < 0.01$), but this relationship was classed only as moderate at 2 min ($r = -0.35$; $P > 0.05$).

Moderate relationships were observed between Yo-Yo IE2 test performance and %HR$_{\text{max}}$ after 1 min ($r = -0.37$; $P > 0.05$) and 2 min of recovery ($r = -0.38$; $P > 0.05$).

Yo-Yo IE2 test performance in relation to competitive level
Differences ($P < 0.01$) were observed in the Yo-Yo IE2 test between elite senior ($n = 92$) and youth female players ($n = 42$) during the start of season period (1774 ± 532 vs 1490 ± 447 m; ES = 0.6). Furthermore, elite senior female populations performed better ($P < 0.01$) in the Yo-Yo IE2 test than domestic ($n = 46$) and sub-elite ($n = 19$) female players (1261 ± 449 and 994 ± 373 m; ES range of 1.0–1.5). Elite youth female players also outperformed ($P < 0.01$) the sub-elite female population (1490 ± 447 vs 994 ± 373 m; ES = 1.2) but this was not evident for domestic female players (Fig. 3).

Positional differences in Yo-Yo IE2 test performance
For elite senior female players ($n = 92$), it was observed that wide midfielders (2057 ± 550 m; $n = 19$) had a higher Yo-Yo IE2 test performance ($P < 0.05$) than central defenders (1588 ± 534 m; $n = 21$) and attackers (1516 ± 401 m; $n = 16$; ES range of 0.9–1.1), but not central midfielders (1764 ± 473 m; $n = 21$) or full-backs (1964 ± 522 m; $n = 15$; ES range of 0.2–0.6; Fig 4). For elite youth players ($n = 42$), differences ($P < 0.05$) were observed between wide (1760 ± 395 m; $n = 6$) and central midfielders (1727 ± 349 m; $n = 12$) vs central defenders (1151 ± 261 m; $n = 9$) but not attackers (1273 ± 161 m; $n = 6$) or full-backs (1480 ± 610 m; $n = 9$; ES range of 0.5–1.8). Interestingly, full-backs demonstrated the highest intra-positional variability.
Yo-Yo IE2 testing in female soccer players

Fig. 3. Differences in Yo-Yo IE2 test performance in various female soccer populations. Data includes elite senior players \((n = 92)\), elite youth players \((n = 42)\), domestic players \((n = 46)\) and sub-elite players \((n = 19)\). **Significantly lower Yo-Yo IE2 test performance than elite senior players \((P < 0.01)\). °Significantly lower Yo-Yo IE2 test performance than elite youth players \((P < 0.01)\).

Fig. 4. Positional differences in Yo-Yo IE2 test performances in elite senior female players \((n = 92)\). Data includes central defenders \((CD)\) \((n = 21)\), full-backs \((FB)\) \((n = 15)\), central midfielders \((CM)\) \((n = 21)\), wide midfielders \((WM)\) \((n = 19)\), attackers \((ATT)\) \((n = 16)\). *Significantly higher Yo-Yo IE2 test performance than CD and ATT \((P < 0.05)\).

Fig. 5. Longitudinal intra-seasonal trends in Yo-Yo IE2 test performance in elite youth female players \((n = 11)\). Preparation training camp \((September)\) vs training camps in the first \((March)\), middle \((May)\) and just prior to the U-20 International Federation of Association Football World Cup finals \((June)\). **Significantly higher Yo-Yo IE2 test performances than preparation training camp \((P < 0.01)\).

Intra-seasonal changes in Yo-Yo IE2 test performance

For a group of 10 elite senior female players, Yo-Yo IE2 test performance remained unchanged across the middle stage of the season when training load and match exposure were constant (November: 1660 ± 321 m vs January: 1724 ± 398 m vs March: 1776 ± 356 m; ES range of 0.2–0.3), with no difference in peak HR for the three test rounds (192 ± 7 vs 191 ± 8 vs 189 ± 7 bpm; ES range of 0.1–0.4). However, elite youth female players \((n = 11)\) Yo-Yo IE2 test performances were found to be 13% higher \((P < 0.01)\) during a training camp scheduled in the middle of the season \((1767 ± 539 m)\) compared with the first training camp of the season \((1564 ± 504 m)\); Fig. 5. The Yo-Yo IE2 test performance was 11% higher \((P < 0.01)\) just prior to the FIFA U-20 Women’s World Cup at the end of the season \((1742 ± 503 m)\) compared with the first training camp of the international season \((ES\ range of 0.3–0.4)\). For a group of elite senior female players \((n = 13)\), Yo-Yo IE2 test performance increased by 14% \((P < 0.01)\) after completing just 4 weeks of intense training prior to the FIFA Women’s World Cup \((2049 ± 283 vs 1803 ± 342 m; ES = 0.8)\), with no differences \((P > 0.05)\) in peak HR for the two test rounds \((189 ± 7 vs 188 ± 6 bpm; ES = 0.2)\). A large correlation was observed between the baseline Yo-Yo IE2 test performance before the training camp and the level of improvement after 4 weeks \((r = -0.57; P < 0.05; n = 13)\). The level of improvement was defined as the delta change in Yo-Yo IE2 test performance between baseline and the 4-week post-training period. Domestic players \((n = 10)\) Yo-Yo IE2 test performances were higher \((P < 0.01)\) at the end compared with the start of the season \((1293 ± 580 vs 1472 ± 580 m; ES = 0.3)\).
While \( VO_{2max} \) remained relatively constant \((P > 0.05)\) during the same seasonal periods \((50.9 \pm 4.8 \text{ vs } 52.3 \pm 4.4 \text{ mL/kg/min}; \text{ES} = 0.3)\).

Relationships between Yo-Yo IE2 test performance, match running performance and maximal oxygen uptake

Large correlations were found between Yo-Yo IE2 test performance and total distance \((r = 0.55; P < 0.05; n = 13; \text{Fig. 6(a)})\) and high-intensity running \((> 15 \text{ km/h})\) for elite senior female players during competitive matches \((r = 0.70; P < 0.01; n = 13; \text{Fig. 6(b)})\). Large correlations were also evident between Yo-Yo IE2 test performance and other classifications of high-intensity activities with various speed thresholds: 18–21 km/h \((r = 0.65; P < 0.01; n = 13; \text{Fig. 6(c)})\), 21–23 km/h \((r = 0.60; P < 0.05; n = 13)\) and 25–27 km/h \((r = 0.58; P < 0.05; n = 13)\). The distances covered at a speed threshold of 23–25 km/h \((r = 0.52; P > 0.05; n = 13)\) and in sprinting \((> 27 \text{ km/h}; r = 0.13; P > 0.05; n = 13)\) produced non significant large and small correlations with Yo-Yo IE2 test performance. The relationship between performance in the Yo-Yo IE2 test and the distance covered in low-intensity during a match by elite senior female players was classified as small \((r = 0.22; P > 0.05; n = 13)\). Based on the analysis of domestic players \((n = 19)\), a large correlation was obtained between Yo-Yo IE2 test performance and \( VO_{2max} \) \((r = 0.68; P < 0.01; \text{Fig. 6(d)})\).

Discussion

The present data demonstrate that the Yo-Yo IE2 test performance is reproducible and test scores are closely related to the running performance of female soccer players during competitive matches. Moreover, the Yo-Yo IE2 test was found to demarcate between performances of female soccer players in different levels of competition and positions. Finally, the Yo-Yo IE2 test seems to be a sensitive tool that is able to monitor and detect seasonal variation in physical capacity of elite

![Figure 6](image-url)
female soccer populations after optimal short-term training interventions.

Test–retest reproducibility of the Yo-Yo IE2 test

Previous studies have evaluated the test–retest reproducibility of various laboratory- and field-based endurance tests in elite male soccer players (Krustrup et al., 2003, 2006b; Kalapotharakos et al., 2011), but limited research has been reported on female soccer populations. We found that domestic female players’ performance in the Yo-Yo IE2 test was similar when repeated within a 7-day period, displaying a CV value of 4.5%. This test–retest value was similar to that obtained for elite male players using the Yo-Yo IE2 test (3.9%, Bradley et al., 2011), but lower than the values observed for the Yo-Yo IR2 test (8.3–9.6%, Michalsik & Bangsbo, 1995; Krustrup et al., 2006b; Bangsbo et al., 2008). The deviation in the CV values between the Yo-Yo IE2 and Yo-Yo IR2 tests could be attributed to the sampled population (Hopkins et al., 2001), the two- to threefold higher test score in the Yo-Yo IE2 test and/or differences in shuttle speeds and recovery periods in each of the tests and the resultant impact on aerobic and anaerobic systems. Furthermore, the reproducibility of the Yo-Yo IE2 test for male and female soccer populations is also in agreement with laboratory endurance tests that elicit VO$_{2}$max and various lactate threshold indices (Weltman et al., 1990; Pitzinger & Freedson, 1998; Dickhuth et al., 1999; Grant et al., 2002; Kalapotharakos et al., 2011). This is the first study to quantify the test–retest reproducibility of female players using the Yo-Yo IE2 test and indicates that they are comparable with their male counterparts. This is not surprising given the recent advances in training and testing methods employed in the female game that allow players to become familiar with exerting themselves maximally. This is a particularly important finding as research indicates gender and training status are influential factors regarding the reproducibility of performance tests (Hopkins et al., 2001). Furthermore, the Yo-Yo IE2 test performance of domestic female players displayed an absolute TE of 67 m, which indicates low test–retest variability and more importantly highlights that a performance deviation of 1.5–2.0 times the TE (100–133 m) would be classified as a worthwhile change in performance (Hopkins, 2000). Based on the reproducibility criteria set out by the research literature for physical tests (Currell & Jeukendrup, 2008; Hopkins et al., 2009), it seems apparent that the Yo-Yo IE2 test exhibits robust reproducibility in domestic female soccer players and should be able to detect small worthwhile changes in performance. It is important to note that the reproducibility of Yo-Yo IE2 test performance in elite female players remains unknown. However, when the domestic female players are split into ‘Low’ and ‘High’ performers (‘Low’ < 1400, n = 16 and ‘High’ > 1400 m, n = 11), the ‘High’ domestic female players demonstrated a very low test–retest CV (2.5%). Thus, based on this sub-analysis of the test–retest data, the Yo-Yo IE2 test appears to be highly reproducible for players with high physical capacity and this could possibly apply to elite female players.

Concurrent and construct validity of the Yo-Yo IE2 test

For the Yo-Yo IE2 test to be classified as a valid tool to assess the intermittent endurance capacity of elite female soccer populations, it must be judged based on various models of validity (logical, criterion, predictive, convergent). Although, the current study can only focus on concurrent and construct models of validity (Currell & Jeukendrup, 2008), more research is needed to establish other forms of validity. Regarding concurrent validity, we observed a large correlation between Yo-Yo IE2 test performance vs total distance ($r = 0.55, P < 0.05$) and high-intensity running ($r = 0.70, P < 0.01$) for elite senior female players during competitive matches. It is worth mentioning that the magnitudes of the correlation coefficients are large based on the classifications of Hopkins et al. (2009). The distance covered by high-intensity running has been proposed to be of great importance for performance in elite female soccer because this parameter clearly distinguishes between performance levels and training status (Krustrup & Bangsbo, 2001; Krustrup et al., 2005; Mohr et al., 2008; Andersson et al., 2010). The Yo-Yo IR1 test has also been shown to be closely related to high-intensity match running performance in both elite male and female soccer players (Krustrup et al., 2003, 2005). However, other authors have also reported correlations of similar magnitude between match running performance and various fitness parameters using physical performance tests, which are less soccer specific (Rampinini et al., 2007; Buchheit et al., 2010). Recently, Bradley et al. (2011) found large to very large correlations between the Yo-Yo IE2 test vs total and high-intensity distances covered in a match for elite male players. In comparison with the relationships obtained in the present study, we found a reduced correlation coefficient for total distance covered, but improved correlation for high-intensity running parameters. Some authors suggest that these large to very large correlation coefficients are not sufficiently strong enough for the Yo-Yo IE2 test to be deemed valid (Mendez-Villanueva & Buchheit, 2011). However, given the complexity of soccer performance and the high match-to-match variability exhibited by elite players (Gregson et al., 2010), the correlation coefficients presented should be placed in context. For instance, relationships of this magnitude have greater significance in complex sports such as soccer, compared with sports in which the criteria performance measure is highly reproducible (e.g., distance running). The trend observed in the present study for the higher correlation between physical capacity and high-intensity running compared...
with total distance further highlights the importance of this variable given its ability to demarcate between different competitive levels in female players with total distance covered similar between levels (Mohr et al., 2008). High-intensity running has also been shown to increase by 23% after training interventions, while total distance remained unaltered (Krustrup & Bangsbo, 2001). Thus, our data suggest that the Yo-Yo IE2 test has good concurrent validity and can be classified as a useful tool to assess the physical capacity of elite female soccer populations. Moreover, this data further highlight the importance of high-intensity running variables as important indicators of match running performance. The present study also revealed a large correlation between $\dot{V}O_{2\text{max}}$ and Yo-Yo IE2 test performance in domestic female players, which could lead to the suggestion that the latter could be used to predict the former (Bangsbo et al., 2008). However, the relationship between the Yo-Yo IE2 test and other variables ($\dot{V}O_{2\text{max}}$ and match running performance) are not sufficiently large ($r > 0.80$) to have predictive ability. This suggests that the estimation of $\dot{V}O_{2\text{max}}$ from the Yo-Yo IE2 test performance cannot be accurately calculated from the current data set because of large residual errors. It is important to point out that based on the relationship between $\dot{V}O_{2\text{max}}$ and Yo-Yo IE2 test performance, the latter could be used as an indicator, but not predictor of the former. The correlations between these measures for female players are higher than those recently obtained from elite and sub-elite male soccer populations (Bradley et al., 2011). Krustrup et al. (2005) proposed that $\dot{V}O_{2\text{max}}$ is more important for the match running performance in elite female players than in elite male soccer. Thus, our data seems to support this contention and could be attributed to the lower anaerobic fitness parameters of female players compared with their male counterparts (Mujika et al., 2009).

Mohr et al. (2008) demonstrated that high-intensity running is a distinguishing characteristic between female players at different competitive levels, whereby elite players perform more high-intensity running than players in lower levels of competition. Thus, to establish construct validity, the Yo-Yo IE2 test should be able to discriminate between various competitive levels in female soccer. The present study found that elite senior female players Yo-Yo IE2 test performances were better than elite youth, domestic and sub-elite players (ES was classified as moderate to large). This finding is in agreement with a plethora of research on male and female soccer populations using the Yo-Yo IR1 and 2 tests (Krustrup et al., 2003, 2005, 2006b; Bangsbo et al., 2008). For instance, Yo-Yo IR1 test performances of elite senior female players were 18% and 38% higher than domestic and sub-elite players, respectively (Bangsbo et al., 2008). Interestingly, larger deviations were observed in the present study between similar competitive standards using the Yo-Yo IE2 test. A possible reason for this finding could be attributed to the caliber of players in the elite senior female group as they were derived from two national sides in the top 12 of the FIFA women’s world rankings and top European teams competing in the UEFA Champion’s League. In support of this, Krustrup et al. (2006a) observed that Yo-Yo IE2 test performances were also higher for female players in the national team compared with those playing in the best domestic Danish league.

Because of the large sample size used within the present study (excluding the sub-elite female population), normative ranges for different female soccer populations can be reported (20–80th percentiles). Sub-elite and domestic female players produced Yo-Yo IE2 test performances of 720–1360 and 920–1520 m, respectively. While elite youth and senior players Yo-Yo IE2 test performances were in the region of 1040–1880 and 1240–2280 m, respectively. Although no data currently exist on Yo-Yo IE2 test performance of sub-elite female players, the elite senior and domestic players in the present study produced similar performances (~1300 and 1700 m) to Danish players of a similar standard (Krustrup et al., 2006a, 2010). All Yo-Yo IE2 test performances observed in the present study for various soccer populations are substantially higher than those reported by Bangsbo et al. (2010) for age-matched untrained women (~240 m). When comparing across gender (Bradley et al., 2011), elite senior female players Yo-Yo IE2 performances are ~30% lower than their male counterparts (~1700 vs 2500 m). The magnitude of the gender difference is very similar to data reported for the Yo-Yo IR1 test in some studies (Krustrup et al., 2003, 2005; Bangsbo et al., 2008), but not others (Mujika et al., 2009). However, the differential in Yo-Yo IE2 test performance between gender is much wider than that reported for world record times in a variety of endurance events (Sparling et al., 1998; Cheuvront et al., 2005) and supports the current assertion that elite female players seem to possess a lower physical capacity than male players across a range of fitness attributes (Rhodes & Mosher, 1992; Tamer et al., 1997; Krustrup et al., 2005, 2010; Stolen et al., 2005; Mujika et al., 2009). Thus, given the relationship between physical capacity and match running performance, it is not surprising that high-intensity running distances covered in elite female matches are lower than their male counterparts of a similar competitive level and age (Krustrup et al., 2005; Mohr et al., 2008; Andersson et al., 2010). This was supported by the current data set as high-intensity running was found to be 30% lower (~1700 m) than that covered by elite male players (~2400 m; Mohr et al., 2003) using an identical speed threshold (> 15 km/h). Interestingly, total distance covered was comparable between elite female players in the present study (~10 600 m) and elite male players in the research literature (~10 700–10 900 m; Mohr et al., 2003; Bradley et al., 2009). One potential limitation of the present study is the use of
pre-defined high-intensity speed thresholds for elite senior female players that are generally applied to elite senior male players. Given that top speed characteristics and anaerobic threshold indices of female players are substantially lower than their male counterparts (Krustrup et al., 2005, 2010; Stolen et al., 2005; Sayers et al., 2008; Abt & Lovell, 2009), speed thresholds should ideally be individualized based on these parameters. For instance, research has indicated that the running speeds at a blood lactate concentration corresponding to 4 mmol/L are 9–10 and 13–14 km/h for elite female and male soccer players, respectively (McMillan et al., 2005; Ingebrigtsen et al., 2011; Kalapotharakos et al., 2011). Using the traditional high-intensity running threshold (> 15 km/h), our data trends are in agreement with studies that have analyzed elite senior female players of a similar competitive standard (1500–1700 m; Mohr et al., 2008; Andersson et al., 2010). Despite the similarities with previous research, it must be noted that the sample of players used in the present study for time motion analysis was small (n = 13), and data was only tracked across a two-match period. Thus, more research is warranted to elucidate the gender differences between elite male and female players using individualized speed thresholds. Based on our findings, it seems apparent that the Yo-Yo IE2 test displays good construct validity, given its ability to discriminate among female players at different levels of performance and vary markedly with male players, but more research using a larger sample needs to be undertaken to fully establish concurrent validity.

### Intra-seasonal changes in Yo-Yo IE2 test performance

For a group of elite youth female players, it was observed that Yo-Yo IE2 test performance was 13% and 11% higher during training camps scheduled in the middle and end of the season compared with the first training camp of the international season; however, this was classified as a small effect (ES = 0.3–0.4). Changes of a moderate magnitude (ES = 0.8) were also found for a group of elite senior female players. The Yo-Yo IE2 test performances of these players improved by 14% after completing only 4 weeks of intense training prior to the FIFA Women’s World Cup (2049 ± 283 vs 1803 ± 342 m), while Yo-Yo IE2 test performances remained unchanged across the most constant period of the season in elite female senior players when no specific interventions were provided. These data are supported by a plethora of research on elite male soccer players using the Yo-Yo IR1 and 2 tests (Krustrup et al., 2003, 2006b; Bangsbo et al., 2008), which report large improvements (25–42%) between the preparation, start, and middle phases of the season when specific training interventions have been introduced during these periods. Previous research also illustrates improvement across the same periods of the season in elite youth male players using the Yo-Yo IE2 test (Bradley et al., 2011), but this was generally of a lower magnitude (10% and 15%) than that reported for the Yo-Yo IR tests. This is the first study to...
report long- and short-term seasonal variation in elite youth and senior female players using the Yo-Yo IE2 test. The trends observed in the present study for elite youth female players are similar to the progressive improvements observed in Yo-Yo IR1 test performance in national female players across a 12-month period (Bangsbo et al., 2008). These trends can be attributed to the focus on fitness training during key stages of the season to achieve peak fitness prior to major tournaments such as the UEFA Women’s European Championships and both the FIFA senior and U-20 Women’s World Cup finals. Although, the ES classifications for intra-seasonal changes varied from a low to moderate (ES range from 0.3 to 0.8), it must be acknowledged that the lower ES magnitudes observed in the present study are not surprising given the elite senior sample of players we assessed and their high level of fitness throughout all periods of the season. This study also observed increases in Yo-Yo IE2 test performance at the end compared with the start of the season (1293 ± 580 vs 1472 ± 580 m) but VO$_{2\text{max}}$ remained relatively constant (50.9 ± 4.8 vs 52.3 ± 4.4 mL/kg/min). Although, it must be stated that the ES calculations were identical for both Yo-Yo IE2 test performance and VO$_{2\text{max}}$ (ES = 0.3). Some authors have failed to observe a significant training-induced change in VO$_{2\text{max}}$ after intense intermittent training despite marked improvement in physical match running performance and Yo-Yo IR test scores (Krustrup & Bangsbo, 2001; Bangsbo et al., 2008). Others have observed large increases in Yo-Yo IR test performance during the preparation period with only small changes in VO$_{2\text{max}}$ (Bangsbo et al., 2006). Taken together, it appears that the Yo-Yo IE2 test is a sensitive tool that can differentiate between intermittent exercise performances of elite female players in various stages of the season, particularly before and after intense training interventions.

Limitations

The reader should be aware of the limitations of the present study. Although stringent guidelines were followed by the authors to standardize the procedures of the Yo-Yo IE2 test, marginal variations will have been evident in the indoor surfaces, environment, and the degree of recovery from previous training sessions and matches as a result of the large sample of players from diverse soccer populations (sub-elite to international players). The low number of players and matches used to determine the complex relationship between physical capacity and match running performance highlights the need for more research to fully elucidate the Yo-Yo IE2 tests concurrent validity. Furthermore, given that the pre-defined high-intensity speed thresholds are identical between the genders then the relationships obtained between the Yo-Yo IE2 test performance and match running performance must be considered carefully. Finally, although we found intra-seasonal variation in Yo-Yo IE2 test performance, no standard model was used to assess players (e.g., the number of weeks and training/match composition would be different from club to club and one seasonal period to the next).

Conclusion

The data clearly demonstrate that the Yo-Yo IE2 test is reproducible and can be used as an indicator of the match-specific physical capacity of female soccer players. Furthermore, the Yo-Yo IE2 test illustrates high sensitivity by differentiating performance of female players in various competitive levels, stages of the season and playing positions.

Perspectives

More studies investigating the physiological demands and physical capacity of female soccer athletes using a standard battery of tests are warranted. The present study incorporates the physical capacity of the largest sample of elite female players published to date and illustrates the importance of using soccer-specific field testing to profile physical performance and to monitoring intra-seasonal variation in fitness levels to establish if training programming can be optimized. Practical importance in the present study must be placed on the application of the Yo-Yo IE2 test to female soccer players. The sensitivity of the Yo-Yo IE2 test has now been quantified and applied sports scientists could potentially use this modality to track female players of sub-elite to elite standard across the season. This would enable them to monitor training progression to allow optimization of training to occur. Finally, because of the relationships observed between the Yo-Yo IE2 test and match running performance, it is clear that this test can be used to gauge training status and match-specific physical performance in elite female soccer players.

Key words: female soccer, physical capacity, match running performance, playing position, competitive level, heart rate.

Acknowledgements

The authors would like to acknowledge the players and their soccer clubs/football associations for their participation. We would also like to thank the National Team Coach Kenneth Heiner-Møller and the Danish Football Association (Dansk Boldspil-Union) for their cooperation. Additional acknowledgement must be given to Michael Mulhern for providing access to the domestic female teams. Thanks must be given to Michele Di Mascio and Daniel Peart for assistance during domestic female testing sessions. The excellent technical assistance by Anja Heiner-Møller, Tine Cedervkist Visker, Stuart Dixon, and Rachael Dawe is appreciated. Team Denmark and the University of Sunderland supported the study. There is no conflict of interest for any of the authors named in this manuscript.
References


